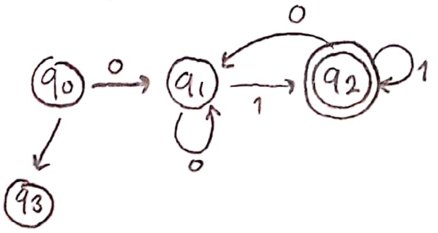
Mesut Şafak Bilici
17011086
Computation Theory | Homework 2

Şafak B.

1. For each of the following languages, state the class of the language (whether it is regular, context-free (but not regular) or neither). Prove your answer. Make sure, if you claim that a language is context free, that you show that is also not regular.

(a) $L = \{w \in \{0,1\}^* : \exists k \geq 0$ and w is a binary encoding (leading zeros allowed) of $2k+1\}$.

There is a DFA for L.



So, the language L is accepted by DFA and thus, L is regular language.

(b) $L = \{a^* b^* c^* - \{a^n b^n c^n : n \geq 0\}\}$.

If L is regular than $\bar{L}$ must be regular.

$\bar{L} = \{w \in \{a,b,c\}^*$ with letters out of order$\} \cup \{a^m b^m c^m : n \geq 0\}$

if $\bar{L}$ is regular than $L_1 = \bar{L} \cap (a^* b^* c^*)$ must be regular.

Then $L_1 = \{a^m b^m c^m ; n \geq 0\}$, which is not context free.

Due to Regular Languages is subset of CFL, L is not regular language.

(c) $L = \{ x \in \{a,b\}^* : |x|$ is even and the first half of x has one more than the second half $\}$.

Let us say L is CFL

and $w = \underbrace{aaa}_{u}\underbrace{aa}_{v}\underbrace{a}_{x}\underbrace{ba}_{y}\underbrace{aa}_{z}$

if L is CFL, uv^ixy^iz must be $\in L$ for $i \geq 0$.

But for $i=2$, new string is $uv^2xy^2z = \underbrace{aaaaaaabaabaa}_{\notin L}$

So L is not CFL.

② $\boxed{\text{Let } L = \{ w \in \{a,b\}^* : \text{the first, middle, and last characters of } w \text{ are identical} \}}$

(a) Show a context-free grammar that generates L.

$S \to aXa \mid bYb$

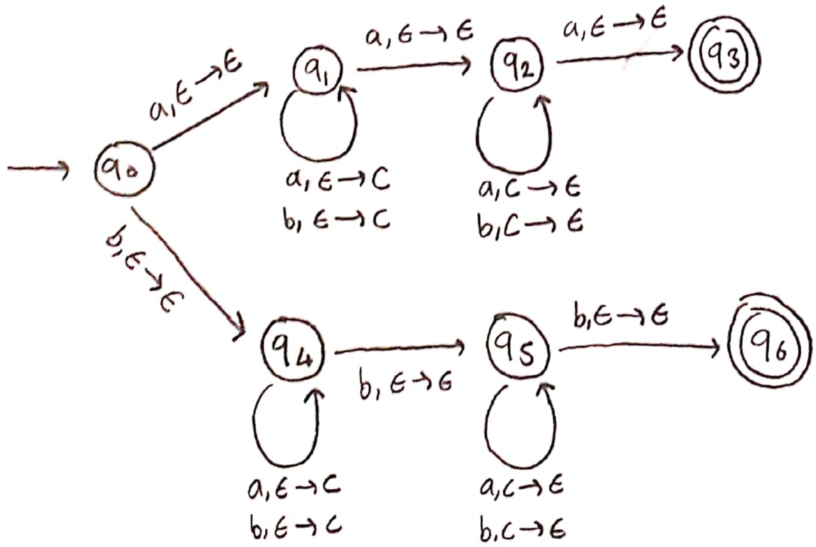
$X \to Y \mid a$

$Y \to X \mid b$

Example:

$aXa \to aaXaa \to$

(b)



(c) Prove that L is not regular.

Let us say that L is regular.

Choose $w = ab^m ab^m a$

m is critical length

$|xy| \leq m$

$|y| \geq 1$

and $y = b^k$

Rewrite...

$$w = ab^m ab^m a = xyz = \underbrace{ab\ldots b}_{x}\underbrace{b\ldots b}_{y}\underbrace{b\ldots b\,a\,\overbrace{b\ldots b}^{m}a}_{z}$$

(m over $b\ldots b\,b\ldots b\,b\ldots b$)

from the pumping lemma

$xy^iz \in L$

Thus: $xy^2z \in L$, $xy^2z = \underbrace{ab\ldots b}_{x}\underbrace{b\ldots b}_{y}\underbrace{b\ldots b}_{y}\underbrace{b\ldots b\,a\,\overbrace{b\ldots b}^{m}a}_{z}$

($m+k$ over the first run of b's)

$= ab^{m+k}ab^m a$

But $ab^{m+k}ab^m a \notin L$ due to middle element is not 'a'.

③ $\boxed{\text{Consider the following grammar } G.}$

$S \to 1S1 \mid T$

$T \to 1X1 \mid X$

$X \to 0X0 \mid 1$

(a) What are the first (shortest) four strings $L(G)$?

- $S \to T \to X \to 1$
- $S \to T \to X \to 0X0 \to 010$
- $S \to T \to 1X1 \to 111$
- $S \to T \to X \to 0X0 \to 00X00 \to 00100$

(b) Give an example of string $w \in \{0,1\}^*$ such that $|w| > 7$ and $w \notin L(G)$

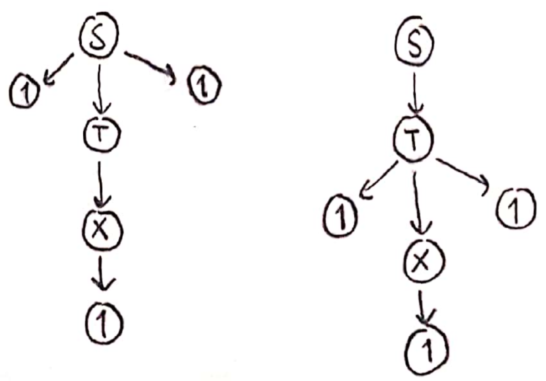
$w = 00010001 \in L(G)$

(c) Show that G is ambiguous.

the string $w = 111$

has two different

derivation tree $\longrightarrow$



4 Find grammars for given languages

(a) $L_1 = L(aaa^*b + b)$

Grammar is:

$S \to aaTb \mid \epsilon$
$T \to aT \mid \epsilon$

b) $L_2 = \{a^n c^m b^n : m, n \geq 1\}$

Grammar is:

$S \to aTb$
$T \to aXb \mid X$
$X \to cY$
$Y \to c \mid \epsilon$

(c) $L_3 = L\{a^n b^{n+1} : n \geq 0\}$

Grammar is:

$S \to aTbb \mid b$
$T \to aTb \mid \epsilon$

5 Convert the grammars given below to Chomsky Normal Form of the. Do not forget to give modified 4-Tuple Grammar.

(a) $S \to a \mid aA \mid B \mid C$
$A \to aB \mid \epsilon$
$B \to Aa$
$C \to bCD$
$D \to bbb$

introducing new variables T_a T_b:

$S \to a \mid T_a A \mid AT_a \mid T_b CD$
$A \to T_a B$
$B \to AT_a \mid a$
$C \to T_b CD$
$D \to T_b T_b T_b$
$T_a \to a$
$T_b \to b$

$\longrightarrow$

introducing intermediate variables V_1, V_2:

$S \to a \mid T_a A \mid AT_a \mid V_1 D$
$A \to T_a B$
$B \to AT_a \mid a$
$C \to V_1 D$
$D \to V_2 T_b$
$T_a \to a$
$T_b \to b$
$V_1 \to T_b C$
$V_2 \to T_b T_b$

CFG

Then:

$G = (\{T_a, T_b, A, B, C, D, V_1, V_2\}, \{a, b\}, S, P)$

Where P is, $P = \{ S \to a \mid T_a A \mid AT_a \mid V_1 D,$
$A \to T_a B,$
$B \to AT_a \mid a,$
$C \to V_1 D,$
$D \to V_2 T_b,$
$V_1 \to T_b C,$
$V_2 \to T_b T_b,$
$T_a \to a,$
$T_b \to b \}$

(b)

$S \to aA \mid aBB$
$A \to aaA \mid \epsilon$
$B \to bB \mid bbC$
$C \to B$

introducing new variables for the terminals T_a, T_b:

$S \to T_a A \mid T_a BB$
$A \to T_a T_a A \mid \epsilon$
$B \to T_b B \mid T_b T_b C$
$C \to B$

$\longrightarrow$

introducing intermediate variables V_1, V_2, V_3:

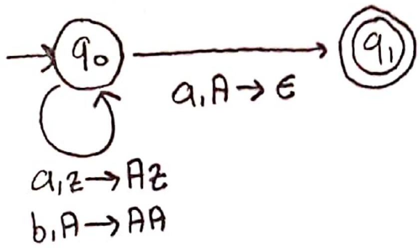
$S \to T_a A \mid T_a V_1 \mid a$
$A \to V_2 A \mid T_a T_a$
$B \to T_b B \mid V_3 B$
$V_1 \to BB$
$V_2 \to T_a T_a$
$V_3 \to T_b T_b$
$T_a \to a$
$T_b \to b$

Then:

$G = (\{T_a, T_b, A, B, V_1, V_2, V_3\}, \{a, b\}, S, P)$

Where $P = \{ S \to T_a A \mid T_a V_1 \mid a,$
$A \to V_2 A \mid T_a T_a,$
$B \to T_b B \mid V_3 B,$
$V_1 \to BB,$
$V_2 \to T_a T_a,$
$V_3 \to T_b T_b,$
$T_a \to a,$
$T_b \to b \}$

⑥ Find a context-free grammar that generates the language accepted by the PDA
$M = (\{q_0, q_1\}, \{a, b\}, \{A, z\}, \delta, q_0, \{q_1\})$, with transitions
$\delta(q_0, a, z) = \{(q_0, Az)\}$
$\delta(q_0, b, A) = \{(q_0, AA)\}$
$\delta(q_0, a, A) = \{(q_1, \epsilon)\}$



$S \to aAb$
$A \to b \mid \epsilon$

⑦ Determine whether the following languages are context-free or not.

(a) $L = \{a^n w w^R a^n : n \geq 0, w \in \{a,b\}^*\}$

There is a grammar for language L:

$S \rightarrow aSa \mid bSb \mid \epsilon$

So, it is context-free.

(b) $L = \{a^n b^j a^n b^j : n \geq 0, j \geq 0\}$

take $z = a^m b^m a^m b^m$, break z into $uvwxy$, where $|vwx| \leq m$ and $vx \neq \epsilon$

Let say v in first a^m and y in first b^m, that means: $v = a^{\eta}$, $y = b^{\theta}$

If we pump up v and y the string will be $a^{m+\eta} b^{m+\theta} a^m b^m$

$a^{m+\eta} b^{m+\theta} a^m b^m \notin L$

Hence L is not not context-free.

(c) $L = \{a^n b^j a^j b^n : n \geq 0, j \geq 0\}$

There is a grammar for language L:

$S \rightarrow bSa \mid T$

$T \rightarrow aTb \mid \epsilon$

Hence it is context-free.

(d) $L = \{a^n b^j a^k b^l : n + j \leq k + l\}$

A language that involves counting or comparsion of three or more variables independently is not context free language.

(e) $L = \{a^n b^j a^k b^l : n \leq k, j \leq l\}$

A language that involves counting or comparsion of three or more variables independently is not context free language.

(f) $L = \{a^n b^n c^j : n \leq j\}$

take $z = a^m b^m c^m$ and break z into $uvwxy$, where $|vwx| \leq m$ and $vx \neq \epsilon$

Let say $v = a^{\eta}$ and $x = a^{\theta}$, pump up, string will be

$uv^0wx^0y = a^{m-\eta-\theta} b^m c^m \notin L$

Hence L is not context-free.

(g) $L = \{w \in L((a+b+c)^*) : n_a(w) = n_b(w) = 2n_c(w)\}$, where $n_x(w)$ denotes the number of symbol X in the string w.

Not context free.

(8) Let L be a context-free language. Proove that there is an integer $p \geq 2$, such that the following is true:

For every string s in L with $|s| \geq p$, there is a string z in L such that $|s| < |z| \leq |s| + p$

?

(9) For the context-free languages in problem 7, find the grammar of the language in Chomsky Normal Form.

(7.a)

introducing new new variables T_a, T_b:

$S \to T_a S T_a \mid T_b S T_b \mid \epsilon$

$T_a \to a$

$T_b \to b$

$\longrightarrow$

introducing intermediate variables V_1, V_2:

$S \to V_1 T_a \mid V_2 T_b \mid T_a T_a \mid T_b T_b$

$T_a \to a$

$T_b \to b$

$V_1 \to T_a S$

$V_2 \to T_b S$

(7.c)

introducing new Variables T_a, T_b:

$S \to T_b S T_a \mid T$

$T \to T_a T T_b \mid \epsilon$

$T_a \to a$

$T_b \to b$

introducing intermediate variables V_1, V_2

$S \to V_1 T_a \mid V_2 T_b \mid T_a T_b$

$T \to V_2 T_b \mid T_a T_b$

$T_a \to a$

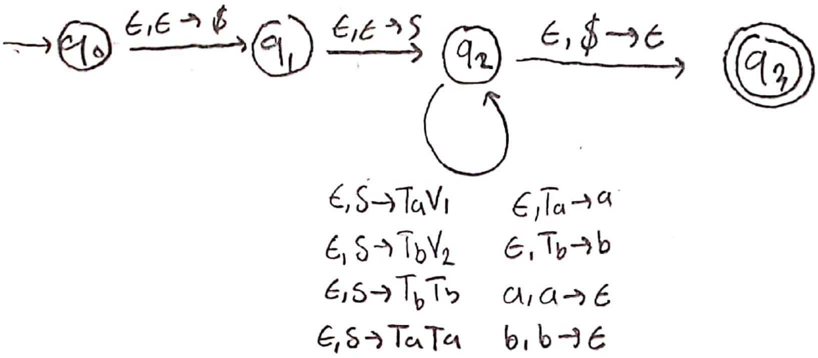
$T_b \to b$

$V_1 \to T_b S$

$V_2 \to T_a T$

10 | For the grammars in problem 9, obtain the PDA using the obtained context free grammars

(9.7.a)



(9.7.c)